Journal of Computational and Applied Mathematics 47 (1993) 277–283
North-Holland

CAM 1306

On the discretization of generalized Riccati differential equations with constant coefficients

Lothar Berg
Fachbereich Mathematik, Universität Rostock, Germany

Received 20 April 1991
Revised 29 April 1992

Abstract

Berg, L., On the discretization of generalized Riccati differential equations with constant coefficients, Journal of Computational and Applied Mathematics 47 (1993) 277–283.

Systems of linear and quadratic differential equations with constant coefficients are approximated by means of a modified trapezoidal rule. In particular, we consider cases where a quadratic or cubic integral exists which remains completely or approximately invariant. The paper ends with an explicit solution of a nonlinear recursion.

Keywords: Riccati equations; trapezoidal rule; invariant integrals.

1. Introduction

Let us consider the real system of quadratic differential equations

$$(z^i)' = z^{\mathrm{T}} A_i z + b_i z + f^i(t), \tag{1}$$

with $i = 1, \dots, m$ for the unknown functions $z^i(t)$ and $z = (z^1, \dots, z^m)^{\mathrm{T}}$. The functions $f^i(t)$ are given and also the constant entries of the $m \times m$ matrices A_i and $B = (b_1, \dots, b_m)^{\mathrm{T}}$ with the rows $b_i = (b_{i1}, \dots, b_{im})$. Without loss of generality we assume the matrices A_i to be symmetric. The system (1) is a generalization of the Riccati system

$$Z' = ZAZ + BZ + ZC + F,$$

Correspondence to: L. Berg, Fachbereich Mathematik, Universität Rostock, Rostock, D-18051, Germany.

with $m \times m$ matrices A, B, C, $F(t)$, $Z(t)$. Choosing a mesh size h and the notations $t_n = nh$, $f_n^i = f^i(t_n)$ with integers n, we replace (1) by the difference equations

$$\frac{1}{h}\left(z_{n+1}^i - z_n^i\right) = z_n^{\mathrm{T}} A_i z_{n+1} + \tfrac{1}{2} b_i \left(z_{n+1} + z_n\right) + \tfrac{1}{2}\left(f_{n+1}^i + f_n^i\right), \tag{2}$$

for the approximations $z_n = (z_n^1, \ldots, z_n^m)$ of $z(t_n)$. In view of the approximations

$$z^{\mathrm{T}}(t) A z(t+h) = z^{\mathrm{T}}(t) A z(t) + z^{\mathrm{T}}(t) A z'(t) h + \mathrm{O}(h^2),$$

$$\tfrac{1}{2}\left[z^{\mathrm{T}}(t+h) A z(t+h) + z^{\mathrm{T}}(t) A z(t)\right] = z^{\mathrm{T}}(t) A z(t) + z^{\mathrm{T}}(t) A z'(t) h + \mathrm{O}(h^2),$$

for $h \to 0$ and any matrix A, the discretization method (2) can be considered as a modified trapezoidal rule. However, contrary to the usual implicit case, system (2) is linear with respect to z_{n+1} and therefore easier to handle.

The trapezoidal rule has the well-known property that various conservation principles of classical mechanics remain valid, in particular some energy integrals remain invariant, cf. [1,3] and the conservative schemes in [6].

2. Linear systems

Preliminarily, we consider a special case of (1), namely the real homogeneous m-dimensional system

$$z' = Bz, \tag{3}$$

and ask for the existence of an integral of the form

$$z^{\mathrm{T}} A z = \text{const.}, \tag{4}$$

where A is a symmetric nonzero $m \times m$ matrix with constant entries.

Theorem 1. *The linear system* (3) *has the integral* (4), *if and only if there exists a constant matrix* C *with* $C^{\mathrm{T}} = -C$ *and*

$$AB = ACA. \tag{5}$$

Proof. (i) The general solution of (3) is

$$z = \mathrm{e}^{Bt} z_0,$$

so that we obtain

$$z^{\mathrm{T}} A z = z_0 \mathrm{e}^{B^{\mathrm{T}} t} A \mathrm{e}^{Bt} z_0.$$

The condition (5) implies $AB^n = (AC)^n A$. Hence, using the antisymmetry of C, we obtain

$$A \mathrm{e}^{Bt} = \mathrm{e}^{ACt} A = \mathrm{e}^{-AC^{\mathrm{T}} t} A = \mathrm{e}^{-B^{\mathrm{T}} t} A,$$

and therefore $z^{\mathrm{T}} A z = z_0^{\mathrm{T}} A z_0$, i.e., the wanted equation (4).

(ii) Conversely, by differentiation we find from (4), in view of (3),

$$z^{\mathrm{T}} A z' = z^{\mathrm{T}} A B z = 0.$$

Since the vector z can be arbitrary, the matrix of the quadratic form must be antisymmetric, so that we obtain

$$AB + B^{\mathrm{T}}A = 0. \tag{6}$$

If A is a regular matrix, then (6) implies (5) with the antisymmetric matrix $C = -A^{-1}B^{\mathrm{T}}$, and (5) can be simplified to $B = CA$. Otherwise, we consider the Jordan normal form

$$A = U^{\mathrm{T}}DU,$$

where U is an orthogonal and D a diagonal matrix, and (6) changes into

$$DS + S^{\mathrm{T}}D = 0,$$

where $S = UBU^{\mathrm{T}}$. With the block representations

$$D = \begin{pmatrix} D_0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad S = \begin{pmatrix} S_0 & S_1 \\ S_2 & S_3 \end{pmatrix},$$

where D_0 is a regular diagonal matrix, we obtain $D_0S_0 + S_0^{\mathrm{T}}D_0 = 0$ and $S_1 = 0$. From this it follows that $S_0 = C_0D_0$ with $C_0 = -D_0^{-1}S_0^{\mathrm{T}}$ and $C_0^{\mathrm{T}} = -C_0$, so that $B = U^{\mathrm{T}}SU$ implies

$$B = U^{\mathrm{T}}\begin{pmatrix} C_0D_0 & 0 \\ S_2 & S_3 \end{pmatrix}U.$$

Finally, we obtain

$$AB = U^{\mathrm{T}}\begin{pmatrix} D_0C_0D_0 & 0 \\ 0 & 0 \end{pmatrix}U = AU^{\mathrm{T}}\begin{pmatrix} C_0 & 0 \\ 0 & 0 \end{pmatrix}UA,$$

and therefore (5), where C is the matrix on the right-hand side between the two factors A. □

Remark 2. If B is a singular matrix and a^{T} a solution of $a^{\mathrm{T}}B = 0$, then $A = aa^{\mathrm{T}}$ satisfies (5) with $C = 0$, and (4) can be simplified to $a^{\mathrm{T}}z = \text{const}$. If A is a regular matrix, then (6) implies $\operatorname{tr} B = 0$.

Theorem 3. *Under the hypotheses of Theorem* 1 *the solutions of the trapezoidal rule*

$$z_{n+1} - z_n = \tfrac{1}{2}hB(z_{n+1} + z_n) \tag{2'}$$

also satisfy (4).

Proof. From (2′) we obtain, by means of (6),

$$A(z_{n+1} - z_n) = -\tfrac{1}{2}hB^{\mathrm{T}}A(z_{n+1} + z_n),$$

and therefore

$$\begin{aligned}(z_{n+1}^{\mathrm{T}} + z_n^{\mathrm{T}})A(z_{n+1} - z_n) &= -\tfrac{1}{2}h(z_{n+1}^{\mathrm{T}} + z_n^{\mathrm{T}})B^{\mathrm{T}}A(z_{n+1} + z_n) \\ &= -(z_{n+1}^{\mathrm{T}} - z_n^{\mathrm{T}})A(z_{n+1} + z_n).\end{aligned}$$

Since a quadratic form is a scalar, we have $z_{n+1}^{\mathrm{T}}Az_n = z_n^{\mathrm{T}}Az_{n+1}$. Hence we obtain

$$z_{n+1}^{\mathrm{T}}Az_{n+1} = z_n^{\mathrm{T}}Az_n,$$

and finally

$$z_n^{\mathrm{T}} A z_n = z_0^{\mathrm{T}} A z_0. \qquad \square$$

Remark 4. The statement of Theorem 3 can be wrong if the integral is not a quadratic form. This is shown by the simple two-dimensional example

$$x' = x, \qquad y' = ay,$$

with the integral $yx^{-a} = \text{const}$. The solutions of the trapezoidal rule

$$x_{n+1} - x_n = k(x_{n+1} + x_n), \qquad y_{n+1} - y_n = ak(y_{n+1} + y_n),$$

with $k = \frac{1}{2}h$ satisfy $y_n x_n^{-b} = \text{const.}$, with

$$b = \left(\ln\frac{1+ak}{1-ak}\right) \Big/ \left(\ln\frac{1+k}{1-k}\right),$$

so that $b \neq a$ for $a \neq 0, 1, -1$, but $b \to a$ for $k \to 0$.

3. Quadratic systems

Now, we pass on to quadratic differential equations. For the simple example $y' = y^2$ the method (2) becomes $y_{n+1} - y_n = hy_n y_{n+1}$, which implies $1/y_{n+1} - 1/y_n = -h$, and therefore $1/y_n - 1/y_0 = -hn$, or, finally, $y_n = y_0/(1 - hny_0)$. This is again the exact solution of the corresponding differential equation in the discrete points, cf. [4].

In the two-dimensional case we have the following result.

Theorem 5. *The quadratic system*

$$x' = bx^2 + 2cxy + 3dy^2, \qquad y' = -3ax^2 - 2bxy - cy^2 \tag{7}$$

has the integral

$$ax^3 + bx^2y + cxy^2 + dy^3 = \text{const.}, \tag{8}$$

and the solutions of the corresponding method (2)

$$\begin{aligned} x_{n+1} - x_n &= h(bx_n x_{n+1} + cy_n x_{n+1} + cx_n y_{n+1} + 3dy_n y_{n+1}), \\ y_{n+1} - y_n &= -h(3ax_n x_{n+1} + by_n x_{n+1} + bx_n y_{n+1} + cy_n y_{n+1}) \end{aligned} \tag{9}$$

satisfy

$$\frac{ax_n^3 + bx_n^2 y_n + cx_n y_n^2 + dy_n^3}{1 + h^2\left[(3ac - b^2)x_n^2 + (9ad - bc)x_n y_n + (3bd - c^2)y_n^2\right]} = \text{const.} \tag{10}$$

Proof. The first statement can easily be checked. To prove the second statement, we first solve the linear system (9) and find

$$\begin{aligned} x_{n+1} &= \frac{1}{\Delta}\left[x_n + h\left(bx_n^2 + 2cx_n y_n + 3dy_n^2\right)\right], \\ y_{n+1} &= \frac{1}{\Delta}\left[y_n - h\left(3ax_n^2 + 2bx_n y_n + cy_n^2\right)\right] \end{aligned} \tag{11}$$

with

$$\Delta = 1 + h^2\left[(3ac - b^2)x_n^2 + (9ad - bc)x_n y_n + (3bd - c^2)y_n^2\right].$$

Now we have to show that under the replacement (11) the equation (10) is independent of n. This was checked on a Philips P-3230 computer by means of the DERIVE 2.01 system. Here we only show the calculations in the special case $b = c = 0$, where

$$\frac{ax^3 + dy^3}{1 + 9h^2adxy} = \text{const.} \tag{10'}$$

and

$$u = \frac{1}{\Delta}(x + 3dhy^2), \qquad v = \frac{1}{\Delta}(y - 3ahx^2), \tag{11'}$$

with $\Delta = 1 + 9adh^2xy$ in simplified notations. From

$$au^3 + bv^3 = \frac{1}{\Delta^3}(ax^3 + dy^3)\left[1 + 27adh^2(dhy^3 + xy - ahx^3)\right]$$

and

$$1 + 9adh^2uv = \frac{1}{\Delta^2}\left[1 + 18adh^2xy + 9adh^2(xy + 3dhy^3 - 3ahx^3)\right],$$

we immediately see the invariance of (10′) under the substitution $x \to u$, $y \to v$ with (11′), and the theorem is proved. □

Let us mention that the inverse of the transformation (11′) has the same form, only with $-h$ instead of h.

Example 6. Another example for quadratic equations is the Volterra–Lotka system

$$x' = x(1 - y), \qquad y' = y(x - 1),$$

with the integral

$$x - \ln x + y - \ln y = \text{const.}, \tag{12}$$

cf. [2]. The method (2)

$$x_{n+1} - x_n = k(x_{n+1} + x_n - y_n x_{n+1} - x_n y_{n+1}),$$
$$y_{n+1} - y_n = k(x_n y_{n+1} + y_n x_{n+1} - y_{n+1} - y_n),$$

with $k = \frac{1}{2}h$ reads explicitly

$$\begin{aligned} x_{n+1} &= \frac{x_n}{\Delta}\left[(1+k)^2 - k(1-k)y_n - k(1+k)x_n\right], \\ y_{n+1} &= \frac{y_n}{\Delta}\left[(1-k)^2 + k(1+k)x_n + k(1-k)y_n\right], \end{aligned} \tag{13}$$

where $\Delta = 1 - k^2 - k(1-k)x_n + k(1+k)y_n$.

Starting with $x_0 = 0.5$ and $y_0 = 1$, computer calculations give for $k = 0.1$ a good approximation of the curve (12) in the (x, y)-plane, cf. Fig. 1, which deforms for increasing k up to

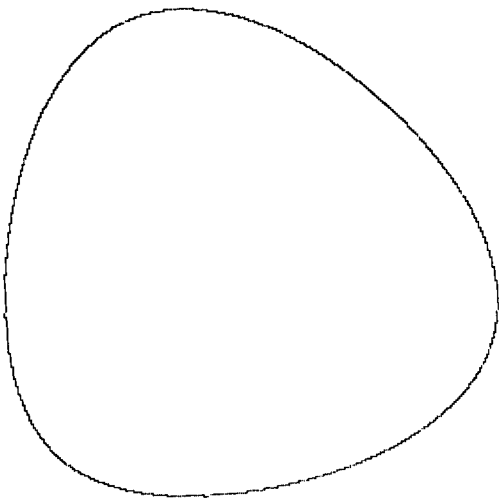

Fig. 1.

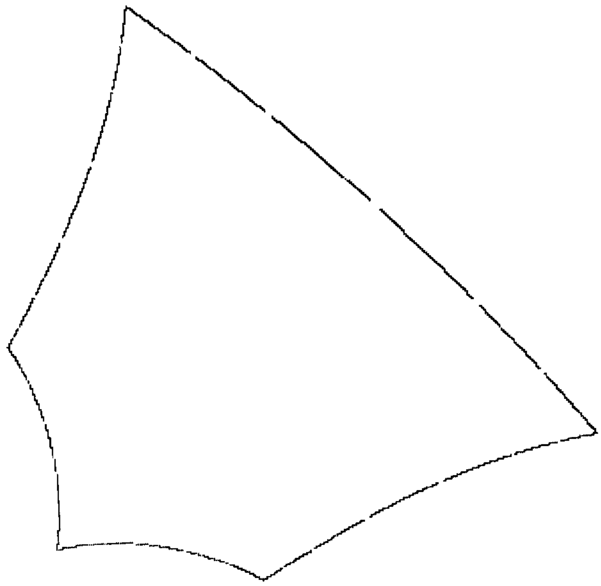

Fig. 2.

$k = 0.675\,69$ to Fig. 2, whereas for $k = 0.6757$ the calculations break down in view of overflow. The curve in Fig. 2 reminds of the curve in [5, Fig. 52].

For $k = 1$ we can construct an explicit solution, because the equations (13) turn over into $x_n = y_{n+1}$ and

$$y_{n+1} y_{n-1} = y_n(2 - y_n).$$

If we introduce the substitution $y_n = 2^{-n} u_n u_{n-1}$, we obtain the new recursion

$$u_n u_{n-3} + u_{n-2} u_{n-1} = 2^n. \tag{14}$$

From $u_{-1} = u_0 = u_1 = 1$ we get the old initial values $x_0 = \frac{1}{2}$, $y_0 = 1$, and we find for the next 23 values u_n with $n \geqslant 2$ the integers

3, 5, 1, 9, 11, 29, −7, 65, 51, 181, −79, 441, 283, 1165,
−599, 2929, 1731, 7589, −4127, 19305, 11051, 49661, −27559,

as well as $u_{-2} = 1$, $u_{-3} = 0$. It is easy to check for these values the linear equations

$$u_{4n} + u_{4n-1} = 2u_{4n-2}, \qquad u_{4n+1} + u_{4n} = 2u_{4n-1},$$
$$u_{4n+2} + u_{4n+1} = 4u_{4n-1}, \qquad u_{4n+3} + u_{4n+2} = 8u_{4n-1},$$

which can be reduced to

$$u_{4n} = 2p_n - q_n, \quad u_{4n+1} = -2p_n + 3q_n, \qquad p_{n+1} = 2p_n + q_n, \quad q_{n+1} = -2p_n + 7q_n \tag{15}$$

with $p_n = u_{4n-2}$, $q_n = u_{4n-1}$. In what follows we need the notations

$$A = \begin{pmatrix} 2 & 1 \\ -2 & 7 \end{pmatrix}, \qquad B = \frac{1}{4}\begin{pmatrix} -4 & 5 \\ 5 & -2 \end{pmatrix}.$$

Then in view of $p_0 = q_0 = 1$ the last two equations of (15) have the explicit solution $(p_n \ q_n)^{\mathrm{T}} = A^n(1 \ 1)^{\mathrm{T}}$. We have to show that (15) with this result solves (14) for all positive integers n. For $4n$ instead of n the left-hand side of (14) becomes, if we drop the index n,

$$\tfrac{1}{2}(2p-q)(-p+q)+pq=\tfrac{1}{2}(-2p^2+5pq-q^2)=(p \quad q)B(p \quad q)^{\mathrm{T}}.$$

Hence, in view of $(1 \ 1)B(1 \ 1)^{\mathrm{T}} = 1$ and $A^{\mathrm{T}}BA = 16B$, which implies $(A^{\mathrm{T}})^n BA^n = 2^{4n}B$, (14) is proved for $n \equiv 0 \bmod 4$. In the remaining three cases with $4n+i$ instead of n, the left-hand side of (14) becomes for $i = 1, 2, 3$,

$$(-2p+3q)p+q(2p-q)=-2p^2+5pq-q^2,$$

$$(2p+q)q+(2p-q)(-2p+3q)=-4p^2+10pq-2q^2,$$

$$(-2p+7q)(2p-q)+(-2p+3q)(2p+q)=-8p^2+20pq-4q^2,$$

so that in any case the right-hand side gets a factor 2, and (14) is proved completely.

The matrix A has the eigenvalues $\lambda_{1,2} = \tfrac{1}{2}(9 \pm \sqrt{17})$, so that y_n is unbounded, and in view of the periodicity of the signs, (x_n, y_n) diverges in form of a spiral.

References

[1] L. Berg, *Differenzengleichungen Zweiter Ordnung mit Anwendungen* (Steinkopff, Darmstadt, 1980).

[2] J. Grasman and E.J.M. Veling, Asymptotic methods for the Volterra–Lotka equations, in: F. Verhulst, Ed., *Asymptotic Analysis,* Lecture Notes in Math. **711** (Springer, Berlin, 1979) 146–157.

[3] D. Greenspan, *Discrete Models* (Addison-Wesley, London, 1973).

[4] R.E. Mickens, Mathematical modelling of differential equations by difference equations, in: D. Lee, R.L. Sternberg and M.H. Schultz, Eds., *Computational Acoustics* (North-Holland, Amsterdam, 1988) 387–393.

[5] H.-O. Peitgen and P.H. Richter, *The Beauty of Fractals* (Springer, Berlin, 1986).

[6] A.A. Samarskij, Numerical methods for problems of mechanics and physics in several dimensions, in: A.A. Samarskij and I. Kátai, Eds., *Mathematical Models in Physics and Chemistry and Numerical Methods of their Realization,* Teubner-Texte Math. **61** (Teubner, Leipzig, 1984) 6–78.